

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2001, 10:03:31 ; Search time 1813.94 Seconds

(without alignments)
11341.049 Million cell updates/sec

Title: US-09-718-754A-1

Perfect score: 1247
Sequence: 1 atcgacataaagttgagatc.....agagtgctagcactagca 1247

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenEmbl:**
1: gb_da:**
2: gb_hcg:**
3: gb_in:**
4: gb_ov:**
5: gb_ov:**
6: gb_ov:**
7: gb_ov:**
8: gb_ov:**
9: gb_ov:**
10: gb_ov:**
11: gb_ov:**
12: gb_ov:**
13: gb_ov:**
14: gb_ov:**
15: gb_ov:**
16: gb_ov:**
17: gb_ov:**
18: gb_ov:**
19: gb_ov:**
20: gb_ov:**
21: gb_ov:**
22: gb_ov:**
23: gb_ov:**
24: gb_ov:**
25: gb_ov:**
26: gb_ov:**
27: gb_ov:**
28: gb_ov:**
29: gb_ov:**
30: gb_ov:**
31: gb_ov:**
32: gb_ov:**
33: gb_ov:**
34: gb_ov:**
35: gb_ov:**
36: gb_ov:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| C 1 | 108 | 8.7 | 6914 | 8 | ZMA251453 |
| C 2 | 102.4 | 8.2 | 9285 | 8 | ZM009989 |
| C 3 | 99 | 7.9 | 4695 | 8 | A69563 |
| C 4 | 97 | 7.8 | 151802 | 8 | ZM009989 |
| C 5 | 97 | 7.8 | 151802 | 8 | ZM009989 |
| C 6 | 70.8 | 5.7 | 151802 | 8 | ZM009989 |
| C 7 | 62.6 | 5.0 | 13868 | 8 | ZM009989 |
| C 8 | 62.6 | 5.0 | 13868 | 8 | ZM009989 |
| C 9 | 59 | 4.7 | 175914 | 17 | ZM009989 |
| C 10 | 58.4 | 4.7 | 321003 | 2 | ZM009989 |
| C 11 | 57.8 | 4.6 | 854 | 8 | ZM009989 |
| C 12 | 57.8 | 4.6 | 1241 | 8 | ZM009989 |
| C 13 | 57.2 | 4.6 | 105574 | 2 | ZM009989 |
| C 14 | 57.2 | 4.6 | 170237 | 9 | ZM009989 |
| C 15 | 56.8 | 4.6 | 1986 | 8 | ZM009989 |
| C 16 | 56.4 | 4.5 | 67970 | 3 | ZM009989 |
| C 17 | 56.4 | 4.5 | 169546 | 2 | ZM009989 |
| C 18 | 55.6 | 4.5 | 54345 | 3 | ZM009989 |
| C 19 | 55.6 | 4.5 | 15643 | 3 | ZM009989 |
| C 20 | 55.0 | 4.5 | 15664 | 3 | ZM009989 |
| C 21 | 55.0 | 4.5 | 178719 | 2 | ZM009989 |
| C 22 | 54.4 | 4.4 | 137889 | 9 | ZM009989 |
| C 23 | 54.4 | 4.4 | 137889 | 9 | ZM009989 |
| C 24 | 54.2 | 4.3 | 137889 | 9 | ZM009989 |
| C 25 | 54.2 | 4.3 | 137889 | 9 | ZM009989 |
| C 26 | 54.2 | 4.3 | 137889 | 9 | ZM009989 |
| C 27 | 54.2 | 4.3 | 137889 | 9 | ZM009989 |
| C 28 | 53.8 | 4.3 | 170143 | 2 | ZM009989 |
| C 29 | 53.8 | 4.3 | 170143 | 2 | ZM009989 |
| C 30 | 53.6 | 4.3 | 150236 | 9 | ZM009989 |
| C 31 | 53.4 | 4.3 | 1141 | 6 | ZM009989 |
| C 32 | 52.6 | 4.2 | 178783 | 9 | ZM009989 |
| C 33 | 52.6 | 4.2 | 85779 | 8 | ZM009989 |
| C 34 | 52.6 | 4.2 | 113880 | 3 | ZM009989 |
| C 35 | 52.4 | 4.2 | 143409 | 9 | ZM009989 |
| C 36 | 52.4 | 4.2 | 161230 | 2 | ZM009989 |
| C 37 | 52.4 | 4.2 | 165278 | 2 | ZM009989 |
| C 38 | 52.4 | 4.2 | 331326 | 2 | ZM009989 |
| C 39 | 52.2 | 4.2 | 1141 | 6 | ZM009989 |
| C 40 | 52.2 | 4.2 | 110000 | 2 | ZM009989 |
| C 41 | 52.2 | 4.2 | 183125 | 2 | ZM009989 |
| C 42 | 51.8 | 4.2 | 14331 | 9 | ZM009989 |
| C 43 | 51.8 | 4.2 | 170781 | 9 | ZM009989 |
| C 44 | 51.8 | 4.2 | 182125 | 2 | ZM009989 |
| C 45 | 51.6 | 4.1 | 8291 | 8 | ZM009989 |

ALIGNMENTS

| RESULT | 1 | ZMA251453 | 6914 bp | DNA | PLN | 04-DEC-2000 |
|------------|---|---|-------------|-----|-----|-------------|
| LOCUS | | ZMA251453/c | | | | |
| DEFINITION | | Zea mays see2a gene for putative legumain, exons 1-9. | | | | |
| ACCESSION | | AJ251453.1 | GI:11558851 | | | |
| VERSION | | AJ251453.1 | GI:11558851 | | | |
| KEYWORDS | | legumain; see2a gene. | | | | |
| SOURCE | | Zea mays. | | | | |
| ORGANISM | | Zea mays. | | | | |
| REFERENCE | | Donnison, I.S., Griffiths, C.M., Thomas, A., Hosken, S.E., Bridges, I. and Thomas, H. | | | | |
| AUTHORS | | Donnison, I.S., Griffiths, C.M., Thomas, A., Hosken, S.E., Bridges, I. and Thomas, H. | | | | |
| TITLE | | Characterisation of See2, a senescence enhanced cDNA from maize with homology to legumain | | | | |
| JOURNAL | | Unpublished | | | | |
| REFERENCE | | 2 (bases 1 to 6914) | | | | |
| AUTHORS | | Donnison, I.S. | | | | |

| | | | |
|--|---------------------------|------------|---------------------------|
| /clone_lib="Sau3AI partial of maize genomic DNA into | | | |
| Charon 40" | | | |
| mRNA | 3019. .3225 | exon | 3019. .3225 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=7 | Intron | /number=7 |
| | 3225. .3259 | | 3225. .3259 |
| | /number=7 | | /number=7 |
| gene | 3225. .3844 | exon | 3225. .3844 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=8 | Intron | /number=8 |
| | 3510. .3629 | | 3510. .3629 |
| | /gene="Mha1" | | /gene="Mha1" |
| exon | 3630. .3721 | Intron | 3630. .3721 |
| | /number=9 | | /number=9 |
| | 3722. .3844 | exon | 3722. .3844 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=10 | | /number=10 |
| CDS | 3927. .4031 | Intron | 3927. .4031 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=11 | exon | /number=11 |
| | 4032. .4112 | | 4032. .4112 |
| | /number=11 | | /number=11 |
| exon | 4113. .4257 | exon | 4113. .4257 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=12 | Intron | /number=12 |
| | 4258. .4378 | | 4258. .4378 |
| | /number=12 | | /number=12 |
| Intron | 4378. .4543 | exon | 4378. .4543 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=13 | | /number=13 |
| | 4544. .4636 | Intron | 4544. .4636 |
| | /number=13 | | /number=13 |
| exon | 4637. .4875 | exon | 4637. .4875 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=14 | Intron | /number=14 |
| | 4876. .4963 | | 4876. .4963 |
| | /number=14 | | /number=14 |
| Intron | 4964. .4996 | exon | 4964. .4996 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=15 | Intron | /number=15 |
| | 4997. .5090 | | 4997. .5090 |
| | /number=15 | | /number=15 |
| exon | 5091. .5172 | exon | 5091. .5172 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=16 | Intron | /number=16 |
| | 5173. .5256 | | 5173. .5256 |
| | /number=16 | | /number=16 |
| Intron | 5257. .5417 | exon | 5257. .5417 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=17 | Intron | /number=17 |
| | 5418. .5512 | | 5418. .5512 |
| | /number=17 | | /number=17 |
| exon | 5513. .5689 | exon | 5513. .5689 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=18 | Intron | /number=18 |
| | 5690. .5767 | | 5690. .5767 |
| | /number=18 | | /number=18 |
| Intron | 5768. .5950 | exon | 5768. .5950 |
| | /gene="Mha1" | | /gene="Mha1" |
| | /number=19 | Intron | /number=19 |
| | 5951. .6035 | | 5951. .6035 |
| | /number=19 | | /number=19 |
| exon | 6036. .one-of(6408, 6409) | exon | 6036. .one-of(6408, 6409) |
| | /number=20 | | /number=20 |
| | /evidence=experimental | | /evidence=experimental |
| | 6408 | | 6408 |
| | /evidence=experimental | | /evidence=experimental |
| Intron | 6409 | polyA_site | 6409 |
| | /evidence=experimental | | /evidence=experimental |
| | 2329 a | | 2329 a |
| | 2181 c | | 2181 c |
| | 2086 g | | 2086 g |
| | 2689 t | | 2689 t |


```

primer_bind    /note="Prem2a 5' LTR"  
              2118..151802  
             /note="Prem2a"  
             21164..22119  
             /note="Prem2a gag protein"  
             /codon_start=1  
             /pseudo  
misc_feature   22689..22692  
             /note="Opie2 target site duplication"  
             22693..31017  
             /organism="Zea mays"  
             /db_xref="taxon:4577"  
             /chromosome="9"  
             /map="9S"  
             /transposon="Opie2"  
             /note="retrotransposon"  
LTR            22693..22939  
             /note="Opie2 5' LTR"  
             23942..23959  
             /note="Opie2"  
             23990..25465  
             /gene="Opie2 gag"  
             /product="Opie2 gag protein"  
             23990..25465  
             /gene="Opie2 gag"  
             23990..25465  
             /codon_start=1  
             /product="Opie2 gag protein"  
             /protein_id="AAK73106.1"  
             /db_xref="GI:14719281"  
             /translation="MDPKGCMVNVNDKEKESFVNPDKEKFPNGSSGGHKRKDKRAKKTKT  
RIKEIIVYSDSSSSQDDNDYERKKKTVMNSNPFYSR...POSSNAHLISIPGRPR  
PKEDEIDYGDSKMSHSLFSHPISIELVENGMHFDMSDPFINDOIHKNAOATTVY  
ILASICDREHYAVSNAKNOIMPTLTLSHEGNKRKKKAYKAKADETHLGKEDSPC  
SSSDENEGEATAFPAFKSALPPNHHCCLMARBEKVVSPTDSYASSPESSODEID  
VSCFLPGIDGTATYKINELIDALDKRRLLICEODLLYEHDITFYAOKSLATETKEN  
EWLSEGLSTHSISSIKSINDPLNAKETLASXTNSREHNVCNCCKPNVDGSHH  
LVCTSKINDVAVNAOLKTSKNDPKLYKADAYATVWRHPSYTDGTCFKRAKLISIR  
HVAPSIKEGCRPMNISTOKNHASLYHDBRRISRNVHDRS"  
             25863..29013  
             /gene="Opie2 pol"  
             /pseudo  
             29014(25863..26686;26718..29013)  
             /gene="Opie2 pol"  
             /note="contains frame-shift mutation; Opie2 pol"  
             /codon_start=1  
             /pseudo  
             29799..29811  
             /note="Opie2 polypurine tract"  
             29813..31017  
             /note="Opie2 3' LTR"  
             31018..31022  
             /note="Opie2 target site duplication"  
             32981..33745  
             /product="Rf2 orf3"  
             32981..33745  
             /codon_start=1  
             /product="Rf2 orf3"  
             /protein_id="AAK73107.1"  
             /db_xref="GI:14719282"  
             /translation="MADKTYIVPCDDPWPESTITVEDQLVALVLDLPGSGDPQEPKEM  
LAPRSEDAPPFGYVSTFEFHGRGFMASTRFALAPHYGVGFHNTPNSIGQAQA  
ILPAACEAGFLIDPHDLWTHTLFSAEFPAASTRYKAMVAVASGCTLDISGCARQAG  
IIASLVSNKGNNMFYLRLNDYVLDPFESORVTYVAANMRNCPHARTRRSSNPFSS  
PIRSCEPALLPGPSPLTAGCSHMOSGCCFRK"  
             36113..36120  
             /note="Ins2 8-bp target site repeat"  
             36121..36492  
             /organism="Zea mays"  
             /db_xref="taxon:4577"  
             /chromosome="9"  
             /map="9S"  
repeat_region source
```

```

repeat_region      /insertion_seq="Ins2"
                   36493..36500
                   /note="Ins2 8-bp target site repeat"
CDS                complement(37488..37784)

Query Match       7.8%; Score 97; DB 8; Length 151802;
Best Local Similarity 77.6%; Pred. No. 4,3e-11;
Matches 159; Conservative 0; Mismatches 35; Indels 11; Gaps 3;

QY 296 agattaaagttagtcttggaatgcacataagactaatgtagtaagaataatgttgagg 355
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4740 AGATTATAGACAAATGTTTAATATCACTAGACGTATGTTCAGCTCTAAAATGAAATAA- 4680

QY 356 acattcaaacacccta-----tcattatgatctattttgtaaattgactaata 406
    ||| ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4681 ACATCCAAACAGCGTCAACTAATAGTGCTCAACTATTAAGCAAATAGCTAATAA 4622

QY 407 gttatagtagtatctttaagcagctcttttttcctgcgaattttltagccacatacat 466
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4621 GTTACCTAGCATTTGTTTACCTACGTATTTTCACATCA- TTTTACGCCACTACTAT 4563

QY 467 tagtttagtgatcttaacaatccc 491
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4562 TAGCTTAGTGCAATCAACACCOC 4538

RESULT 6
LOCUS AF391808 151802 bp DNA PLN 16-JUL-2001
DEFINITION Zea mays bz locus, partial sequence.
ACCESSION AF391808
VERSION AF391808.1 GI:14719277
KEYWORDS .
SOURCE Zea mays.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 113967 to 117446)
AUTHORS Ralston,E.J., English,J., and Dooner,H.K.
TITLE Sequence of three bromo alleles of maize and correlation with the genetic fine structure
JOURNAL Genetics. 119 (1), 185-197 (1988)
MEDLINE 88284304
PUBMED 3396861
2 (bases 91198 to 151802)
AUTHORS Fu,H., Park,W., Yan,X., Zheng,Z., Shen,B., and Dooner,H.K.
TITLE From the Cover: The highly recombinogenic bz locus lies in an unusually gene-rich region of the maize genome
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 98 (15), 8903-8908 (2001)
MEDLINE 21352966
PUBMED 11436866
3 (bases 1 to 151802)
REFERENCE Fu,H., Zheng,Z., and Dooner,H.K.
AUTHORS Large differences in recombination rates within adjacent gene-dense and retrotransposon regions of maize
JOURNAL Unpublished
4 (bases 1 to 151802)
REFERENCE Fu,H., Zheng,Z., and Dooner,H.K.
AUTHORS Direct Submissions
JOURNAL Submitted (15-JUN-2001) Waksman Institute, Rutgers University, 190 Frelinghuysen Rd, Piscataway, NJ 08854, USA
FEATURES
source location/Qualifiers
1..151802
/organism="Zea mays"
/db_xref="taxon:4577"
/chromosome="9"
/map="95"
join(21..178,283..478,1451..1927,3754..3942)
/gene="hypr04"
/note="similar to Arabidopsis thaliana hypothetical protein"
```

```

gene      /product="hypothetical protein 4"
21..3942
/gene="hypoc4"
join(21..178,203..478,1451..1927,3754..3942)
CDS      /gene="hypoc4"
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein 4"
/protein_id="AAK73103.1"
/db_xref="GI:14719278"
/translacion="MEVGDLHKVMEVRLAKTKPDATARAFLDRVAROVPIRMRRHM
RVKVLSESPENRPLGLWAGVGVKRLRBRDDEPIYEEVLDITMLHELCHNR
GPHDKOFLKLMDELKRGITGTGSGDGRVGVFVHPPELQATLAAQKRRAN
GALLPSGPRKLGENSEIEMSLSPVOAAMAERWMDLMTCSHOASLIDSDVITL
OESPULTTRDGKDKGCSNWTSPSTSRTHIARDRDRTSDALDSMRNCGACATLN
OPIAPICVCCTTKPKIAKATTTMSCFCTLENTKDKCSACDMKRTSGPVATY
GRTSD"
7823..9043
/mRNA     /gene="mapk1"
/note="similar to MAP kinase kinase"
/product="MAP kinase kinase"
7823..9043
/gene="mapk1"
7823..9043
/gene="mapk1"
/codon_start=1
/product="MAP kinase kinase"
/protein_id="AAK73104.1"
/db_xref="GI:14719279"
/translacion="MSDQALFSNRAPRPHKRAVGRPRITQESRRHHYHPRPNT
LPRPAQSDPRTTQCRARASPPSPARARARISPPTPSNRSPSPSPSLSR
OPPEPGSPRRASPPRRSSSSRPPLRLERVARVSGAGETVMVVRHRTGPR
YALKVLGNHDAVKRIARETALITAEHRAVRCHEMKEGELDLLRYMGSL
DGRITAEPRADVGKVSGLIATLHRHVIYHROIKPSNLLIDSRVKIADREVGIT
LNOTMDPNSVSGIAYNSPERINTDLNDSYDAGDIWSGLSILEYIGRPRE
NLGRGDMAALMCAICYNDPPEPPPTASREFGFIACCLQKNPARLITPAQLQHPV
AGHPQPLAAPPS"
complement(join(12719..13170,13252..14404))
/gene="cyp450"
/product="cytochrome P450"
complement(12719..14404)
/gene="cyp450"
complement(join(12719..13170,13252..14404))
/gene="cyp450"
/codon_start=1
/product="cytochrome P450"
/protein_id="AAK73105.1"
/db_xref="GI:14719280"
/translacion="MGTGVALVAAALAVVYVPTALALHLVMPRYAVARAROCI
KSPFRLVNGKEKVDMMKATSGDTLDLASHDIIRAKMFOIRANVSLIGVFLTMSG
STPFLVSHDMVRVLDKAGLIGRTDPTGLTSLDKRGVFTGDGWSHRRVHP
AFAMKRLKMTGAAACAAEVIRGWEARAAASGDREVEVYQQPTLTDVISHA
FGSSYRQKGEFLAQRELOPIAFASINGMVRPOTQAPKA VVRWQLEKTVNGTLM
AIIIDERLAAAEANGYGSDLGLMLEAANGDGKROQOAMSDEILDECKTEFFACH
DITSHLLTAAWFLGTHPEMOQLREEVYRGGAEPVLSGDALKKLVTMVLVETL
RLYGAVPLIARATAGADLCGVKVPKGTLLIPIAMLRDEEYWGADAGANPREFD
GNGRAATHPNALLSFSIGPRSCIGDPFALMBAKATLALLIRPAFRVAPDEVHADDF
LTIQSGILPYVILLEPAVLVTSSG"
complement(15815..16428)
/gene="gag"
complement(15815..16428)
/gene="gag"
complement(15815..16428)
/codon_start=1
/pseudo  1737..19832
/note="RIR2 gag/pol"
/codon_start=1
/pseudo  19900..19904
/note="Pren2a target site duplication"
19905..100704
/organism="Zea mays"
/db_xref="taxon:4577"

/misc_feature
repeat_region
/chromosome="9"
/map="9S"
/transposon="Pren2a"
/note="retrotransposon"
19905..21116
/note="Pren2a 5'LTR"
21118..151802
/note="Pren2a"
21164..22119
/note="Pren2a gag protein"
/codon_start=1
/pseudo  22688..22692
/note="Opie2 target site duplication"
22693..31017
/organism="Zea mays"
/db_xref="taxon:4577"
/chromosome="9"
/map="9S"
/transposon="Opie2"
/note="retrotransposon"
22693..23939
/note="Opie2 5'LTR"
23942..23959
/note="Opie2"
23990..25465
/gene="Opie2 gag"
/product="Opie2 gag protein"
23990..25465
/gene="Opie2 gag"
23990..25465
/gene="Opie2 gag"
/codon_start=1
/product="Opie2 gag protein"
/protein_id="AAK73106.1"
/db_xref="GI:14719281"
/translacion="MDPKGKGVYNDKRESFVNDPKDEKPTNSGSGHRRKRRKT
RRIKEIYYVDSDESSSSQKDDNDYERKKTVNSNPSFYSKIPQSSNAHLISIPGKR
PHRDEGVGFMSHKMRSHLSHPSTIMELVENGHFDMSDPMTINDQIHKAQAATV
LNASLGRDYKRVSVLONAKOIMDTLKLISHBQNKREKKAKKAKOETHLGRKEMPSDC
SSSDSNBGLATTAENKSALEFPNHHCTCLMARKKVSTYRDSTYASSSDESDIED
YSCILFKGDRYTVKRIKINELIDALDNKRLLEKQEDLLYEBHDFVEAKOSLLEIKRN
EMSCSELSTCHDSISLSKSIINDLNKLEIATKTSREHVMICNRCKDPNDACSEH
LVCSIKLNDENVASLNAOKTSKNDPDKLKFARADATVGRHPTWDGIGFRRKAKMLIS
HKAPISIKKGRKAPMENSIOKNASLYHNDRRYSRVHHDR2"
25863..29013
/pseudo  29013
/gene="Opie2 pol"
join(25863..26686,26718..29013)
/note="Opie2 pol"
/note="contains frame-shift mutation; Opie2 pol"
/codon_start=1
/pseudo  29799..29811
/note="Opie2 polypurine tract"
29813..31017
/note="Opie2 3'LTR"
31018..31022
/note="Opie2 target site duplication"
32981..33745
/product="RIR2 off3"
32981..33745
/codon_start=1
/product="RIR2 off3"
/protein_id="AAK73107.1"
/db_xref="GI:14719282"
/translacion="NADKTIYIPPCDPMFSTIYVEDIOLVAVEDLLHPLSGPPEM
LARSADPTPPGTVSETTEFRBGRGAPASREKRALPHITGEVFNHNPNSIQOA
IFAAVCEGLGIDPIHMDLWTLFSAEFPAASDVAKMAVASGGCTLRSGCARY
ILASLVSNKGMQNMWETFLRNDRVILPEPSQSVVTAAGNNRMGPHAKTRSSNPFSV
PCSSSKPALLPKRGSSPLSTAGCASHMOGCGGCFRR"
36113..36120
/note="Ins2 8-bp target site repeat"

```

| Query Match | Best Local Similarity | Matches 125; Conservative | 5.7%; Score 70.8; DB 8; Length 151802; Pred. No. 2.3e-05; Mismatches 67; Indels 4; Gaps 1; |
|-------------|---|---------------------------|--|
| Query Match | Best Local Similarity | Matches 125; Conservative | 5.7%; Score 70.8; DB 8; Length 151802; Pred. No. 2.3e-05; Mismatches 67; Indels 4; Gaps 1; |
| DB 630 | attctacacgtcgagcgctgctaagaagaacgcgcgtgtcgtctaaagatattacactagcg | 669 | |
| DB 86150 | AAATCTCACCTCGCGGCTTTATTAAGAAACCGCAGTGAATAATGCACTTTTCACTGCG | 86209 | |
| DB 690 | gttggtgagcaaacactgcgtctgaaagaagcagctccaccagagcccttagcttgcacgc | 749 | |
| DB 86210 | GTTCGAAATTAACCGCCCGCGTGAAGATTCGATTCACACTGACCCCTTAACACTG | 86265 | |
| DB 750 | gaacataaaaacgcctcagtaagaaatagctcttagatcgtcactatagactctatagct | 809 | |
| DB 86266 | GTACTGAAAAGCCAGCGATTAATGTTTGAACACCGCAGCTATAGAGCTTCGTGTACT | 86325 | |
| DB 810 | tagtggttagactga 825 | | |
| DB 86326 | AGTGCTCTTAACAAATTA 86341 | | |
| RESULT 7 | AX027357 | 13868 bp | DNA PAT 16-SEP-2000 |
| LOCUS | Sequence | 5 from Patent WO0037488. | |
| DEFINITION | AX027357 | | |
| ACCESSION | AX027357.1 | GI:10188332 | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | Zea mays. | | |
| ORGANISM | Zea mays. | | |
| REFERENCE | Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea. | | |
| AUTHORS | 1 (bases 1 to 13868) | | |
| TITLE | Theissen, G., DeJeu, W., Saedler, H. and Cacharron, J. | | |
| JOURNAL | Novel made-box genes and uses thereof | | |
| FEATURES | Patent: WO 0037488-A; 5 29-JUN-2000; MAX PLANCK GESellschaft (DE) | | |
| source | Location/Qualifiers | | |
| BASE COUNT | 1..13868 | | |
| ORIGIN | /organism="Zea mays" /db_xref="taxon:4577" | | |
| Query Match | Best Local Similarity | Matches 191; Conservative | 5.0%; Score 62.6; DB 6; Length 13868; Pred. No. 0.0012; Mismatches 144; Indels 12; Gaps 3 |
| DB 248 | tgaccttaacataactaaacccaagatatagtagagatgttagatagattaaagrtga | 307 | |
| DB 12085 | TAAATTTAAGAAATTTGACTCAAGATTAATGGAATGTCACATTTTGGAAGAGAGAA | 12144 | |
| DB 308 | tgt-ttgaatgacacagagcctagatgttagtagtcaaatatagttgagacattcaaca | 366 | |
| DB 12145 | GGTCTCAATTGACACATATAGCCATATATTAACTGCAATTAATTTGTAGACATTC | 12204 | |
| DB 367 | cctatacaatttttagtat-----tttagtaaatatagttcaatagtttagt | 416 | |
| DB 12205 | GCTGACCAATATTTCACTATTAAGCTTTTATGTTAAATGCTCAATAGCTTGCACAC | 12264 | |
| DB 417 | tatttaagaactagcttlltlltactagacaattllttagcacaactaaactagtttagt | 476 | |

| | | | | |
|------------|---|--|----------------------------------|-------------|
| Ox | 12265 | TATTGTGTAACCACTAATTCACACAGTTT | TTTTTTTATFAGATAACAA- GCTCTCAGT | 12323 |
| Oy | 477 | gtattcaaatcccccttaagccgctlaagatga | tgccttctcagaactctaaccgatcagtg | 536 |
| Db | 13224 | GCAATTCATAATGAGCCGAGATPATTCTCCA | GCGCGTGATGAGATTTCATTAGTTGGGCTTGG | 12383 |
| Oy | 537 | agacacatttcataagglygtactcglttaagt | caccgcagtagata | 583 |
| Dd | 12384 | GTNGTCGTGGTTGATTGCATCATGAAGTAGT | AACGACGACGCCGCA | 12430 |
| RESULT | 8 | AC020619 | 266371 bp DNA HTG | 17-AUG-2001 |
| LOCUS | AC020619/c | Mus musculus clone RP23-5A1 strain C57BL/6J, | WORKING DRAFT | |
| DEFINITION | | SEQUENCE, 8 unordered pieces. | | |
| ACCESSION | AC020619 | | | |
| VERSION | AC020619.8 | GI:15208565 | | |
| KEYWORDS | HTG; HNGS_PHASE1; HNGS_DRAFT; HNGS_FULLTOP; | HNGS_ACTIVEFIN. | | |
| SOURCE | house mouse. | | | |
| ORGANISM | Mus musculus | | | |
| REFERENCE | Fumakura, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| AUTHORS | Mamukula; Euleria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| TITLE | 1 (bases 1 to 266371) | | | |
| JOURNAL | Genome Therapeutics Corporation Sequencing Center: Mouse Genome | | | |
| REFERENCE | Sequence Data | | | |
| AUTHORS | Unpublished | | | |
| TITLE | 2 (bases 1 to 266371) | | | |
| JOURNAL | Smith,D.R. | | | |
| REFERENCE | Direct Submission | | | |
| AUTHORS | Submitted (06-JAN-2000) Genome Therapeutics Corporation, 100 Beaver | | | |
| TITLE | Street, Waltham, MA 02453, USA | | | |
| JOURNAL | On Aug 17, 2001 this sequence version replaced gi:14010801. | | | |
| COMMENT | ----- Genome Center | | | |
| | Center: Genome Therapeutics Corporation | | | |
| | Center code: GTC | | | |
| | Web site: http://www.genomecorp.com/ | | | |
| | Contact: gtc.sequencet@genomecorp.com | | | |
| | ----- Project Information | | | |
| | Center project name: mg005 | | | |
| | ----- Summary Statistics | | | |
| | Sequencing vector: N/A | | | |
| | Chemistry: dye-terminator Big Dye; 100% of reads | | | |
| | Assembly program: Phrap; version 990315 | | | |
| | Consensus quality: 253434 bases at least Q40 | | | |
| | Consensus quality: 255981 bases at least Q30 | | | |
| | Consensus quality: 257373 bases at least Q20 | | | |
| | Insert size: 265770; sum-of-contigs | | | |
| | Quality coverage: 6.2x In Q20 bases; sum-of-contigs | | | |
| | ----- | | | |
| * | NOTE: This is a 'working draft' sequence. It currently | | | |
| * | consists of 8 contigs. The true order of the pieces | | | |
| * | is not known and their order in this sequence record is | | | |
| * | arbitrary. Gaps between the contigs are represented as | | | |
| * | runs of N, but the exact sizes of the gaps are unknown. | | | |
| * | This record will be updated with the finished sequence | | | |
| * | as soon as it is available and the accession number will | | | |
| * | be preserved. | | | |
| * | 1 | 1229: contig of 1229 bp in length | | |
| * | 1230 | 1329: gap of unknown length | | |
| * | 1303 | 2533: contig of 1204 bp in length | | |
| * | 2534 | 2633: gap of unknown length | | |
| * | 12786 | contig of 10153 bp in length | | |
| * | 12787 | 12886: gap of unknown length | | |
| * | 12887 | contig of 13844 bp in length | | |
| * | 12888 | gap of unknown length | | |
| * | 26730 | contig of 21271 bp in length | | |
| * | 26731 | gap of unknown length | | |
| * | 26830 | gap of unknown length | | |
| * | 26831 | gap of unknown length | | |
| * | 48102 | contig of 30330 bp in length | | |
| * | 48202 | contig of 70718 bp in length | | |
| * | 78531 | gap of unknown length | | |
| * | 78532 | gap of unknown length | | |
| * | 149349: | contig of 70718 bp in length | | |

| | | |
|----|---------------|------------------------------|
| FT | repeat_region | /rpt_family="SINE/Alu" |
| FT | | 7666. .1887 |
| FT | | /rpt_unit=AluY |
| FT | repeat_region | complement(8533. .9044) |
| FT | | /rpt_family="LINE/L1" |
| FT | | /rpt_unit=L1MD3 |
| FT | repeat_region | 9654. .9993 |
| FT | | /rpt_family="SINE/MIR" |
| FT | | /rpt_unit=MIR |
| FT | repeat_region | 10289. .10397 |
| FT | | /rpt_family="LINE/L2" |
| FT | | /rpt_unit=L2 |
| FT | repeat_region | complement(10551. .10861) |
| FT | | /rpt_family="SINE/Alu" |
| FT | | /rpt_unit=AluDb |
| FT | repeat_region | complement(10947. .11117) |
| FT | | /rpt_family="SINE/MIR" |
| FT | | /rpt_unit=MIR |
| FT | repeat_region | 11164. .11232 |
| FT | | /rpt_family="LINE/L2" |
| FT | | /rpt_unit=L2 |
| FT | repeat_region | complement(12045. .12215) |
| FT | | /rpt_family="SINE/MIR" |
| FT | | /rpt_unit=MIR |
| FT | repeat_region | complement(12652. .13133) |
| FT | | /rpt_family="LINE/L1" |
| FT | | /rpt_unit=L1PAl6 |
| FT | repeat_region | complement(13134. .13447) |
| FT | | /rpt_family="SINE/Alu" |
| FT | | /rpt_unit=AluJo |
| FT | repeat_region | complement(13448. .13593) |
| FT | | /rpt_family="LINE/L1" |
| FT | | /rpt_unit=L1PAl6 |
| FT | repeat_region | 13604. .13724 |
| FT | | /rpt_family="LINE/L1" |
| FT | | /rpt_unit=L1ME1 |
| FT | repeat_region | 14467. .14559 |
| FT | | /rpt_family="LINE/Other" |
| FT | | /rpt_unit=HAL1 |
| FT | repeat_region | 14670. .14709 |
| FT | | /rpt_family="Low_complexity" |
| FT | | /rpt_unit=AT-rich |
| FT | repeat_region | 16700. .16843 |
| FT | | /rpt_family="SINE/MIR" |
| FT | | /rpt_unit=MIR |
| FT | repeat_region | complement(17102. .17273) |
| FT | | /rpt_family="SINE/MIR" |
| FT | | /rpt_unit=MIR |
| FT | repeat_region | complement(17502. .17804) |
| FT | | /rpt_family="SINE/Alu" |
| FT | | /rpt_unit=AluY |
| FT | repeat_region | complement(17944. .18048) |
| FT | | /rpt_family="SINE/MIR" |
| FT | | /rpt_unit=MIR |
| FT | repeat_region | 18206. .18514 |
| FT | | /rpt_family="LINE/L2" |
| FT | | /rpt_unit=L2 |
| FT | repeat_region | 18737. .18763 |
| FT | | /rpt_family="Simple_repeat" |
| FT | | /rpt_unit=(A)n |
| FT | repeat_region | 19670. .19945 |
| FT | | /rpt_family="LINE/L1" |
| FT | | /rpt_unit=L1MA7 |
| FT | repeat_region | 20341. .20438 |
| FT | | /rpt_family="LINE/L2" |
| FT | | /rpt_unit=L2 |
| FT | repeat_region | complement(20572. .20598) |
| FT | | /rpt_family="DNA/MER1_type" |
| FT | | /rpt_unit=MER102 |
| FT | repeat_region | complement(20599. .20898) |
| FT | | /rpt_family="SINE/Alu" |
| FT | | /rpt_unit=AluBr |
| FT | repeat_region | complement(20899. .21074) |

| | |
|----|------------------------------|
| FT | /rpt_family="DNA/MER1_type" |
| FT | /rpt_unit="MER102 |
| FT | 22527. .22886 |
| FT | /rpt_family="SINE/Alu" |
| FT | /rpt_unit="Alusx |
| FT | complement(33114. .23225) |
| FT | /rpt_family="LINE/L2" |
| FT | /rpt_unit="L2 |
| FT | complement(23310. .23533) |
| FT | /rpt_family="SINE/MIR" |
| FT | /rpt_unit="MIR |
| FT | 24356. .24669 |
| FT | /rpt_family="SINE/Alu" |
| FT | /rpt_unit="Alusp |
| FT | 25444. .25634 |
| FT | /rpt_family="LINE/L1" |
| FT | /rpt_unit="LIM44 |
| FT | 25652. .25925 |
| FT | /rpt_family="SINE/Alu" |
| FT | /rpt_unit="Alusx |
| FT | 25926. .25958 |
| FT | /rpt_family="Simple_repeat" |
| FT | /rpt_unit="(GGAA)n |
| FT | 25976. .26224 |
| FT | /rpt_family="LINE/L2" |
| FT | /rpt_unit="L2 |
| FT | complement(26225. .26547) |
| FT | /rpt_family="SINE/Alu" |
| FT | /rpt_unit="Aluy |
| FT | 26548. .26703 |
| FT | /rpt_family="LINE/L2" |
| FT | /rpt_unit="L2 |
| FT | complement(26957. .27259) |
| FT | /rpt_family="SINE/Alu" |
| FT | /rpt_unit="Alusx |
| FT | 29841. .30089 |
| FT | /rpt_family="SINE/MIR" |
| FT | /rpt_unit="MIR |
| FT | 30156. .30214 |
| FT | /rpt_family="Low_complexity" |
| FT | /rpt_unit="rich |
| FT | 30243. .30293 |
| FT | /rpt_family="Simple_repeat" |
| FT | /rpt_unit="(ATGGTG)n |
| FT | 30445. .30473 |
| FT | /rpt_family="Simple_repeat" |
| FT | /rpt_unit="(TG)n |
| FT | complement(31164. .31376) |
| FT | /rpt_family="SINE/MIR" |
| FT | /rpt_unit="MIR |
| FT | 31445. .31593 |
| FT | /rpt_family="DNA/MER1_type" |
| FT | /rpt_unit="MER5A |
| FT | complement(31811. .31905) |
| FT | /rpt_family="SINE/MIR" |

| | | | | |
|-----------------------|-----------------|-------------------|-----------|----------------|
| Query Match | 4.7% | Score 59; | DB 17; | Length 175914; |
| Best Local Similarity | 45.1%; | Pred. No. 0.0089; | | |
| Matches 218; | Conservative 0; | Mismatches 265; | Indels 0; | Gaps 0; |

[illegible]

TITLE
JOURNAL
REFERENCE
AUTHORS

Homo sapiens, clone RP11-6D2
2 (bases 1 to 105574)
Britten, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Bateman, J., Brown, A., Castle, A., Colangelo, M., Collins, S., Collins, A., Cooke, P., Dearfield, K., Dewar, K., Donnelly, M., Donnelly, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardy, S., Grant, G., Haggis, B., Hefford, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karas, A., Kiehn, J., Lebeck, J., Liu, C., Locke, K., MacDonald, P., Marcus, N., McKernan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Seaver, P., Stange, P., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Tittell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced g1:5289194.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WITB

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

Project information

Center project name: L2882

Center clone name: 6_D_2

NOTE: This record contains 112 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying relationships that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 876 975: gap of 100 bp
* 976 1849: contig of 874 bp in length
* 1850 1949: gap of 100 bp
* 1950 2811: contig of 862 bp in length
* 2812 2911: gap of 100 bp
* 2912 3742: contig of 831 bp in length
* 3743 3842: gap of 100 bp
* 3843 4717: contig of 875 bp in length
* 4718 4817: gap of 100 bp
* 4818 5413: contig of 596 bp in length
* 5414 5513: gap of 100 bp
* 5514 6398: contig of 885 bp in length
* 6399 6498: gap of 100 bp
* 6499 7384: contig of 886 bp in length
* 7385 7484: gap of 100 bp
* 7485 8357: contig of 873 bp in length
* 8358 8457: gap of 100 bp
* 8458 9306: contig of 849 bp in length
* 9307 9406: gap of 100 bp
* 9407 10257: contig of 851 bp in length
* 10258 10357: gap of 100 bp
* 10358 11213: contig of 856 bp in length
* 11214 11313: gap of 100 bp
* 11314 12162: contig of 849 bp in length
* 12163 12262: gap of 100 bp
* 12263 13138: contig of 876 bp in length
* 13139 13238: gap of 100 bp
* 13239 14081: contig of 843 bp in length

14082 14181: gap of 100 bp
* 14182 15085: contig of 904 bp in length
* 15086 15185: gap of 100 bp
* 15186 16034: contig of 849 bp in length
* 16035 16134: gap of 100 bp
* 16135 16996: contig of 862 bp in length
* 16997 17096: gap of 100 bp
* 17097 17938: contig of 842 bp in length
* 17939 18038: gap of 100 bp
* 18039 18902: contig of 864 bp in length
* 18903 19002: gap of 100 bp
* 19003 19638: contig of 836 bp in length
* 19639 19938: gap of 100 bp
* 19939 20820: contig of 882 bp in length
* 20821 20920: gap of 100 bp
* 20921 21793: contig of 873 bp in length
* 21794 21833: gap of 100 bp
* 21834 22781: contig of 888 bp in length
* 22782 22881: gap of 100 bp
* 22882 23730: contig of 849 bp in length
* 23731 23830: gap of 100 bp
* 23831 24494: contig of 664 bp in length
* 24495 24594: gap of 100 bp
* 24595 25452: contig of 858 bp in length
* 25453 25552: gap of 100 bp
* 25553 26413: contig of 861 bp in length
* 26414 26513: gap of 100 bp
* 26514 27371: contig of 858 bp in length
* 27372 27471: gap of 100 bp
* 27472 28398: contig of 927 bp in length
* 28399 28498: gap of 100 bp
* 28499 29355: contig of 857 bp in length
* 29356 29455: gap of 100 bp
* 29456 30343: contig of 888 bp in length
* 30344 30443: gap of 100 bp
* 30444 31343: contig of 900 bp in length
* 31344 31443: gap of 100 bp
* 31444 32303: contig of 860 bp in length
* 32304 32403: gap of 100 bp
* 32404 33294: contig of 891 bp in length
* 33295 33394: gap of 100 bp
* 33395 34263: contig of 869 bp in length
* 34264 34363: gap of 100 bp
* 34364 35298: contig of 935 bp in length
* 35299 35398: gap of 100 bp
* 35399 36267: contig of 869 bp in length
* 36268 36367: gap of 100 bp
* 36368 37216: contig of 849 bp in length
* 37217 37316: gap of 100 bp
* 37317 38168: contig of 852 bp in length
* 38169 38268: gap of 100 bp
* 38269 38903: contig of 635 bp in length
* 38904 39003: gap of 100 bp
* 39004 39880: contig of 877 bp in length
* 39881 39980: gap of 100 bp
* 39981 40850: contig of 870 bp in length
* 40851 40950: gap of 100 bp
* 40951 41839: contig of 889 bp in length
* 41840 41939: gap of 100 bp
* 41940 42806: contig of 867 bp in length
* 42807 42906: gap of 100 bp
* 42907 43626: contig of 720 bp in length
* 43627 43726: gap of 100 bp
* 43727 44580: contig of 854 bp in length
* 44581 44680: gap of 100 bp
* 44681 45529: contig of 849 bp in length
* 45530 45629: gap of 100 bp
* 45630 46365: contig of 736 bp in length
* 46366 46465: gap of 100 bp
* 46466 47383: contig of 918 bp in length
* 47384 47483: gap of 100 bp
* 47484 48358: contig of 875 bp in length
* 48359 48458: gap of 100 bp

```

      48459 49351: contig of 893 bp in length
      49352 49451: gap of 100 bp
      49452 50334: contig of 883 bp in length
      50335 50434: gap of 100 bp
      50435 51284: contig of 850 bp in length
      51285 51384: gap of 100 bp
      51385 52241: contig of 857 bp in length
      52242 52341: gap of 100 bp
      52342 53192: contig of 851 bp in length
      53193 53292: gap of 100 bp
      53293 54150: contig of 858 bp in length
      54151 54250: gap of 100 bp
      54251 54081: contig of 631 bp in length
      54082 54981: gap of 100 bp
      54982 55844: contig of 863 bp in length
      55845 55944: gap of 100 bp
      55945 56845: contig of 901 bp in length
      56846 56945: gap of 100 bp
      56946 57842: contig of 897 bp in length
      57843 57942: gap of 100 bp
      57943 58819: contig of 877 bp in length
      58820 58919: gap of 100 bp
      58920 59822: contig of 903 bp in length
      59823 59922: gap of 100 bp
      59923 60795: contig of 873 bp in length
      60796 60895: gap of 100 bp
      60896 61764: contig of 869 bp in length
      61765 61864: gap of 100 bp
      61865 62735: contig of 871 bp in length
      62736 62835: gap of 100 bp
      62836 63694: contig of 859 bp in length
      63695 63794: gap of 100 bp
      63795 64660: contig of 866 bp in length
      64661 64760: gap of 100 bp
      64761 65632: contig of 872 bp in length
      65633 65732: gap of 100 bp
      65733 66621: contig of 889 bp in length
      66622 66721: gap of 100 bp
      66722 67617: contig of 896 bp in length
      67618 67717: gap of 100 bp
      67718 68618: contig of 901 bp in length
      68619 68718: gap of 100 bp
      68719 69558: contig of 840 bp in length

```

Query Match 4.6%; Score 57.2; DB 2; Length 105574;
 Best Local Similarity 45.8%; Pred. No. 0.021;
 Matches 220; Conservative 0; Mismatches 259; Indels 1; Gaps 1;

```

OY 5 tacataaagttgattacatacattacactaataagaacacattattaa 64
DB 34659 TACCTTAATATTTATTAATTTACATTAATATTAATAATTAATGTAAT 34718
OY 65 aaactaaacacattacacgtagagagaagacatacattgctattt 124
DB 34719 AAAATTAATAAACTTAATTAATTAATAATAATAATAATAATAATAAT 34778
OY 125 atgcatttaaacacccctcgacacatcagcagtggtgtagcttaactgatac 184
DB 34779 ATATTAATTAATTAATTAATTAATTAATAATAATAATAATAATAATAAT 34838
OY 185 tcttgcttaactactagcaccacatgatatattccttcgacacatgctattat 244
DB 34839 ACATTAATTAATTAATTAATAATAATAATAATAATAATAATAATAATAAT 34898
OY 245 tgttgacttaacatacctaacaagatatagtagagagtgttagtagagata 304
DB 34899 ATATTAATTAATTAATAATAATAATAATAATAATAATAATAATAATAAT 34958
OY 305 tgaatgttaacacacagacataatagtagtagtaaatagtagtagagacata 364
DB 34959 TTAATTTTAATTTATTTTA-TTAATAATTAATTTTATTAATAATAATAAT 35017
OY 365 caaccatacaattatagttatttcttgtagtaaatagtagtaagtagttatta 424

```

```

DB 35018 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 35077
OY 425 agcagcttttttactagcactttttagcacaactagtagtagtagtagtag 484
DB 35078 ATTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 35137

RESULT 14
LOCUS AL137070
DEFINITION Human DNA sequence from clone RP11-251017 on chromosome 9, complete
sequence.
ACCESSION AL137070
VERSION AL137070.9 GI:9367917
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 170237)
REFERENCE
AUTHORS Skuce,C.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
COMMENT
On Jul 22, 2000 this sequence version replaced g1:8977609.
requests: clonerequest@sanger.ac.uk
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu), where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/chr9
RP11-251017 is from the library RPI1-11.1 constructed at the
Roswell Park Cancer Institute by the group of Pletzer de Jong. For
further details see http://bacpac.med.buffalo.edu/
VECTOR: pBac3.6
This sequence is the entire insert of clone RP11-251017.
FEATURES
source
1..170237
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-251017"
/clone_1kb="RPI1-11.1"
/68224..98258
/feature="Random repeat. Single clone region. Assembly
confirmed by restriction enzyme digest data."
98259..98301
/feature="Random repeat. Short join in tandem repeat
confirmed by ECOR1 and HindIII digest data."
98302..98403
/feature="Random repeat. Single clone region. Assembly
confirmed by restriction enzyme digest data."
54028 a 32613 c 32057 g 51539 t
ORIGIN

```


[illegible]

PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
XX a single surface -
PS
Example 6; Page 127; 159pp; English.

CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.

5Q sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

| Query Match | 12.9%; | Score 161.2; | DB 22; | Length 938 |
|-----------------------|--------|--------------|--------|------------|
| Best Local Similarity | 0.99; | Prod No 5 | 6a-30. | |

Matches 7; Conservative 511; Mismatches 254; Indels 0; Gaps 0.

[illegible]

PR 26-JUL-1999; 990S-0145695.
 PR 17-MAR-2000; 2000GS-0190259.

(CLIN-) CLINICAL MICRO SENSORS INC.

XX
 XX
 XX RM;
 PI Umek RM;
 DR WPI: 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface -
 XX
 XX
 XX Example 6; Page 127; 159pp; English.

CC The present invention relates to a composition comprising two nucleic
 CC acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 CC
 XX

SQ Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

| | Query Match | Best local similarity | Score | DB 22; | Length | 936; | | | |
|----|--|-----------------------|-------|------------|--------|--------|----|------|----|
| | Matches | 8; Conservative | 466; | Mismatches | 212; | Indels | 0; | Gaps | 0; |
| Oy | 1 atcgtcaataagcttgatataacattccattacaactaaatagaaccattaat | 60 | : | : | : | : | : | : | : |
| Dd | 96 wncwccc | 155 | : | : | : | : | : | : | : |
| Oy | 61 taaaaaacctaaaacctttaccgcgtaagagagagagcgatcaactgtctatag | 120 | : | : | : | : | : | : | : |
| Dd | 156 wwwwwwccc | 215 | : | : | : | : | : | : | : |
| Oy | 121 ttlttatgatttaaacccttcgaacacatcagcagtggatgatgttcaactgaact | 180 | : | : | : | : | : | : | : |
| Dd | 216 wwwwwwccc | 275 | : | : | : | : | : | : | : |
| Oy | 181 aaatatcgtctttaactagtagcacactgatatatactcttcgaacacgtttat | 240 | : | : | : | : | : | : | : |
| Dd | 276 wwwwwwccc | 335 | : | : | : | : | : | : | : |
| Oy | 241 ccatgttgacttaatacattcaataatccaagatattagtagagcttgatagat | 300 | : | : | : | : | : | : | : |
| Dd | 336 wwwwwwccc | 395 | : | : | : | : | : | : | : |
| Oy | 301 aaggtagatgttgaaagctagagcataatgttagtagctaataatgcttgagacatt | 360 | : | : | : | : | : | : | : |
| Dd | 396 aawwwccc | 455 | : | : | : | : | : | : | : |
| Oy | 361 caaacacccatacatatattatttttagtaattagtagtagttagtatt | 420 | : | : | : | : | : | : | : |
| Dd | 456 wwwwwwccc | 515 | : | : | : | : | : | : | : |
| Oy | 421 tataagctagctttttactagcaattctttagccaactaaatagtttagtgtat | 480 | : | : | : | : | : | : | : |
| Dd | 516 wwwwwwccc | 575 | : | : | : | : | : | : | : |
| Oy | 481 tcaaatacccttagcgcttaagtagtgctcttctgaattcaacogtattgagagac | 540 | : | : | : | : | : | : | : |
| Dd | 576 wwwwwwccc | 635 | : | : | : | : | : | : | : |
| Oy | 541 aacatttcaatagtgtaactggtttagtacccgctcagtgataataatttccatag | 600 | : | : | : | : | : | : | : |
| Dd | 636 wwwwwwccc | 695 | : | : | : | : | : | : | : |
| Oy | 601 gtctcttaagaacacgcgctagctaatgatatattcaactagcggcgtctaaagaaac | 660 | : | : | : | : | : | : | : |
| Dd | 696 wwwwwwccc | 755 | : | : | : | : | : | : | : |

OY 661 cgcccgtaagaatattacacta 686
 DB 756 www. 781

RESULT 10

AAF58259
 ID AAF58259 standard; DNA; 936 BP.

AAF58259;

24-APR-2001 (first entry)

DE Oligonucleotide D2004.

Electron-transfer group; ETM; mismatch; genotyping;
 gene expression; ss.

XX Synthetic.

WO200107665-A2.

01-FEB-2001.

26-JUL-2000; 2000MO-US20476.

26-JUL-1999; 99US-0145695.

17-MAR-2000; 2000US-0190259.

(CLIN-) CLINICAL MICRO SENSORS INC.

Umek RM;

WPI; 2001-159728/16.

Nucleic acids containing electron-transfer group, useful as labels in
 hybridization assays, e.g. for genotyping, allowing repeat analyses on
 a single surface

Example 6; Page 128; 159pp; English.

The present invention relates to a composition comprising two nucleic
 acids each containing an electron-transfer group (ETM) having
 different redox potentials. The invention is used for electronic
 detection of nucleic acids, especially of substitutions (mismatches)
 and single-nucleotide polymorphisms, e.g. for genotyping,
 monitoring gene expression.

Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match 12.9%; Score 160.4; DB 22; Length 936;
 Best Local Similarity 1.2%; Pred. No. 8,7e-30;
 Matches 8; Conservative 466; Mismatches 212; Indels 0; Gaps 0;

OY 1 atcgatataaagttgattatatacatatcattcaactaataagaacacattatc 60
 DB 96 www. 155
 OY 61 taataaactaataaacccttaccgcgtacagagagagacatcaactgtattag 120
 DB 156 www. 215
 OY 121 ttatgcatttaaacaccttcgaacatcagcagtggtatgattagtttaactatc 180
 DB 216 www. 275
 OY 181 aatctctgtcttaactagcaccaactgataataactttcgaaacatgtattat 240
 DB 276 www. 335
 OY 241 ctatgttgacttaactaactaataacaaagatatagtagagatgttagatagatt 300
 DB 300 www. 360

DB 336 www. 395
 OY 301 aagtgatgtttgaatgcactagagctaaatagttagctaaatagttgagacatt 360
 DB 396 www. 455
 OY 361 caaacacccatcaatattagttatttttagtaaatagtaatagttagttatc 420
 DB 456 www. 515
 OY 421 tataagctagcttttttactagcaatttttagccacactaacatlagtttagtatt 480
 DB 516 www. 575
 OY 481 tcaatacccttaagcgttaagtgaagtcctctctagaacttaacgcatgtggagac 540
 DB 576 www. 635
 OY 541 aacatttcatagtgtaactgttaagtcacgcgtcagtgataataatattcacatgcg 600
 DB 636 www. 695
 OY 601 gttccttaagcaaacgcgcagtcagtattacactagcggcgtgctaagaataac 660
 DB 696 www. 755
 OY 661 cgcccgtaagaatattacacta 686
 DB 756 www. 781

RESULT 11

AAF58262
 ID AAF58262 standard; DNA; 936 BP.

AAF58262;

24-APR-2001 (first entry)

DE Oligonucleotide D2007.

Electron-transfer group; ETM; mismatch; genotyping;
 gene expression; ss.

XX Synthetic.

WO200107665-A2.

01-FEB-2001.

26-JUL-2000; 2000MO-US20476.

26-JUL-1999; 99US-0145695.

17-MAR-2000; 2000US-0190259.

(CLIN-) CLINICAL MICRO SENSORS INC.

Umek RM;

WPI; 2001-159728/16.

Nucleic acids containing electron-transfer group, useful as labels in
 hybridization assays, e.g. for genotyping, allowing repeat analyses on
 a single surface

Example 6; Page 128; 159pp; English.

The present invention relates to a composition comprising two nucleic
 acids each containing an electron-transfer group (ETM) having
 different redox potentials. The invention is used for electronic
 detection of nucleic acids, especially of substitutions (mismatches)
 and single-nucleotide polymorphisms, e.g. for genotyping,
 monitoring gene expression.

XX Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;

Query Match 12.9%; Score 160.4; DB 22; Length 936;
Best Local Similarity 1.2%; Pred. No. 8.7e-30;
Matches 8; Conservative 466; Mismatches 212; Indels 0; Gaps 0;

QY 1 atcgatcataaagtgtgattatcaaatccattacaactaaatagaacccaattaat 60
DB 96 www. 155
QY 61 taaaaaactaaaaaaccttatacaccgtacagagagagagacatcgtcattag 120
DB 156 www. 215
QY 121 ttattatgatttaaacaccttcgaacctcagcagcgtgtgttagtttaactgatact 180
DB 216 www. 275
QY 181 aatatctgtctttaactaactagcacaactgataataacttcgaacacatgtattat 240
DB 276 www. 335
QY 241 ctatgttgacttaactaactaactaactaactaactaactaactaactaactaact 300
DB 336 www. 395
QY 301 aagtgatgttgactgactgactgactgactgactgactgactgactgactgact 360
DB 396 www. 455
QY 361 caaacacccatacattatagtatttttgtaaatagttatagttatgtatt 420
DB 456 www. 515
QY 421 tattaagcagcttttttactagcaattttttagcacaactaactagtttttagttat 480
DB 516 www. 575
QY 481 tcaaatacccttaagcgttaagtctcttctctgaatcttaacctgctgagagac 540
DB 576 www. 635
QY 541 aacatttcataagtgactgtttaagtcacgtcagtgagataattttcacatgcg 600
DB 636 www. 695
QY 601 gtctcttaagcaacgcgctgtctaatgatatattacactagcggctgctaaagaaac 660
DB 696 www. 755
QY 661 cgcgcgtgctaaagatatattacacta 686
DB 756 www. 781

RESULT 12

AAFS8255 AAF58255 standard; DNA: 938 BP.

XX AAF58255;

XX 24-APR-2001 (first entry)

DE Oligonucleotide D1876.

XX Ejection-transfer group; ETM; mismatch; genotyping;

XX gene expression; ss.

XX Synthetic.

XX WO200107665-A2.

XX 01-FEB-2001.

PD 26-JUL-2000; 2000MO-US20476.

PF 26-JUL-1999; 9905-0145695.

PR 17-MAR-2000; 2000US-0190259.

XX (CLIN-) CLINICAL MICRO SENSORS INC.

PA UmeK RM;

PI WPI: 2001-159728/16.

XX Nucleic acids containing ejection-transfer group, useful as labels in

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on

PS a single surface

XX Example 6; Page 127; 159pp; English.

XX The present invention relates to a composition comprising two nucleic

CC acids each containing an ejection-transfer group (ETM) having

CC different redox potentials. The invention is used for electronic

CC detection of nucleic acids, especially of substitutions (mismatches)

CC and single-nucleotide polymorphisms, e.g. for genotyping,

XX monitoring gene expression.

SO Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match 12.9%; Score 160.4; DB 22; Length 938;
Best Local Similarity 1.2%; Pred. No. 8.7e-30;
Matches 8; Conservative 466; Mismatches 212; Indels 0; Gaps 0;

QY 1 atcgatcataaagtgtgattatcaaatccattacaactaaatagaacccaattaat 60
DB 96 www. 155
QY 61 taaaaaactaaaaaaccttatacaccgtacagagagagagacatcgtcattag 120
DB 156 www. 215
QY 121 ttattatgatttaaacaccttcgaacctcagcagcgtgtgttagtttaactgatact 180
DB 216 www. 275
QY 181 aatatctgtctttaactaactagcacaactgataataacttcgaacacatgtattat 240
DB 276 www. 335
QY 241 ctatgttgacttaactaactaactaactaactaactaactaactaactaactaact 300
DB 336 www. 395
QY 301 aagtgatgttgactgactgactgactgactgactgactgactgactgactgact 360
DB 396 www. 455
QY 361 caaacacccatacattatagtatttttgtaaatagttatagttatgtatt 420
DB 456 www. 515
QY 421 tattaagcagcttttttactagcaattttttagcacaactaactagtttttagttat 480
DB 516 www. 575
QY 481 tcaaatacccttaagcgttaagtctcttctgaatcttaacgctatgtgagagac 540
DB 576 www. 635
QY 541 aacatttcataagtgactgtttaagtcacgtcagtgagataattttcacatgcg 600
DB 636 www. 695

This sequence represents the sequence of a genomic clone (P19.14) encoding a maize phytase protein. The sequence was isolated from a

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2001, 11:11:57 ; Search time 89.35 Seconds
(without alignments)
3139,723 Million cell updates/sec

Title: US-09-718-754A-1

Perfect score: 1247

Sequence: 1 atcgataaagtgatgtat.....agagtcgtagcaactagca 1247

Scoring table:

IDENTITY NUC
Gapox 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/pdata/2/1na/5A_COMB.seq:*
2: /cgn2_6/pdata/2/1na/5B_COMB.seq:*
3: /cgn2_6/pdata/2/1na/6A_COMB.seq:*
4: /cgn2_6/pdata/2/1na/6B_COMB.seq:*
5: /cgn2_6/pdata/2/1na/PCPTS_COMB.seq:*
6: /cgn2_6/pdata/2/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| C 1 | 102.6 | 8.2 | 4137 | 3 | US-08-750-357-7 |
| C 2 | 43 | 3.4 | 6768 | 1 | US-08-107-755A-1 |
| C 3 | 43 | 3.4 | 8457 | 1 | US-07-991-667B-1 |
| C 4 | 43 | 3.4 | 8457 | 2 | US-08-544-332-1 |
| C 5 | 42.6 | 3.4 | 7015 | 4 | US-09-177-249-6 |
| C 6 | 42 | 3.4 | 4539 | 1 | US-08-119-512-1 |
| C 7 | 42 | 3.4 | 4539 | 1 | US-08-488-015B-1 |
| C 8 | 42 | 3.4 | 4542 | 3 | US-08-814-412-11 |
| C 9 | 41.4 | 3.3 | 1431 | 4 | US-09-316-083-2 |
| C 10 | 41 | 3.3 | 19124 | 2 | US-08-487-825B-13 |
| C 11 | 40.8 | 3.3 | 10729 | 1 | US-07-781-355-1 |
| C 12 | 40.4 | 3.2 | 615 | 4 | US-08-998-416-186 |
| C 13 | 40.4 | 3.2 | 636 | 4 | US-08-998-416-1137 |
| C 14 | 40.4 | 3.2 | 636 | 4 | US-08-998-416-1137 |
| C 15 | 40.4 | 3.2 | 837 | 4 | US-08-998-416-288 |
| C 16 | 40.4 | 3.2 | 1609 | 4 | US-09-377-648-1 |
| C 17 | 39.8 | 3.2 | 2334 | 1 | US-08-062-632-4 |
| C 18 | 39.4 | 3.2 | 3095 | 6 | 5231168-1 |
| C 19 | 38.8 | 3.1 | 828 | 4 | US-08-998-416-538 |
| C 20 | 38.6 | 3.1 | 615 | 4 | US-08-998-416-186 |
| C 21 | 38.6 | 3.1 | 681 | 4 | US-09-299-378-1 |
| C 22 | 38.6 | 3.1 | 3889 | 2 | US-08-648-298-1 |
| C 23 | 38.2 | 3.1 | 837 | 4 | US-08-998-416-288 |
| C 24 | 38 | 3.0 | 3095 | 6 | 5231168-1 |
| C 25 | 37.8 | 3.0 | 724 | 4 | US-08-998-416-683 |
| C 26 | 37.8 | 3.0 | 3528 | 4 | US-08-984-320-2 |
| C 27 | 37.8 | 3.0 | 3528 | 4 | US-08-467-087A-2 |

| | | | | | | |
|------|------|-----|------|---|--------------------|--------------------|
| C 28 | 37.8 | 3.0 | 4383 | 6 | 5175095-4 | Patent No. 5175095 |
| C 29 | 37.8 | 3.0 | 4383 | 6 | 5177307-1 | Patent No. 5177307 |
| C 30 | 37.6 | 3.0 | 854 | 4 | US-08-998-416-514 | Sequence 534, App |
| C 31 | 37.2 | 3.0 | 656 | 4 | US-08-998-416-525 | Sequence 593, App |
| C 32 | 37.2 | 3.0 | 711 | 4 | US-08-998-416-786 | Sequence 786, App |
| C 33 | 37.2 | 3.0 | 732 | 4 | US-08-998-416-1036 | Sequence 1036, App |
| C 34 | 37.2 | 3.0 | 828 | 4 | US-08-998-416-538 | Sequence 538, App |
| C 35 | 37.2 | 3.0 | 2435 | 4 | US-09-306-593-1 | Sequence 1, App11 |
| C 36 | 36.8 | 3.0 | 663 | 4 | US-08-998-416-191 | Sequence 191, App |
| C 37 | 36.8 | 3.0 | 860 | 4 | US-08-998-416-287 | Sequence 287, App |
| C 38 | 36.8 | 3.0 | 5852 | 1 | US-07-867-106-2 | Sequence 32, App1 |
| C 39 | 36.6 | 2.9 | 1850 | 3 | US-08-617-860B-32 | Sequence 32, App1 |
| C 40 | 36.6 | 2.9 | 4098 | 2 | US-08-605-106-4 | Sequence 1036, App |
| C 41 | 36.6 | 2.9 | 6152 | 4 | US-08-973-463-1 | Sequence 4, App11 |
| C 42 | 36.4 | 2.9 | 663 | 4 | US-08-998-416-191 | Sequence 191, App |
| C 43 | 36.4 | 2.9 | 711 | 4 | US-08-998-416-786 | Sequence 786, App |
| C 44 | 36.4 | 2.9 | 724 | 4 | US-08-998-416-683 | Sequence 683, App |
| C 45 | 36.4 | 2.9 | 732 | 4 | US-08-998-416-1036 | Sequence 1036, App |

ALIGNMENTS

RESULT 1
US-08-750-357-7/c
Sequence 7, Application US/08750357
Patent No. 6008437
GENERAL INFORMATION:
APPLICANT: KREBBERS, ENNO
APPLICANT: WILKINS, MARK
TITLE OF INVENTION: USE OF ANTHOCYANIN GENES TO MAINTAIN
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DONE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,357
FILING DATE: 21-MAR-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 018030-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO. 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: EcolR-HindIII region of plasmid pCOL13
FEATURE:
NAME/KEY: prim_transcript
LOCATION: 188
FEATURE:
NAME/KEY: exon
LOCATION: 188..212
FEATURE:

```

NAME/KEY: -
LOCATION: 4132..4137
OTHER INFORMATION: /Label= HindIII
FEATURE:
NAME/KEY: -
LOCATION: 4114..4137
OTHER INFORMATION: /label= polylinker
OTHER INFORMATION: /note= "part of polylinker of pUC19"
US-08-750-357-7

Query Match      8.2%; Score 102.6; DB 3; Length 4137;
Best Local Similarity 72.2%; Pred. No. 3.6e-18;
Matches 151; Conservative 0; Mismatches 49; Indels 9; Gaps 1.

QY   295  tagattaagcggaatttcgaaatgcacacagcagaactaatagtttagtcaaaaattgtctgga 354
      | | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db337  TTGACTTAAGGGCTGTGGTTTACATCATCTGAACCTAATATCATCTGCTAAATAATTGCAGCA 348
QY   355  gacaatcaaacacccta-----ccaactactagttaattttaagtaatttgatcat 405
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db347  GACATCAAAAACAATATGACTAATTAATTCAGCTAATTAATCCTTTTGACAAAATATGACTAAT 348
QY   406  agttgataggatatatttaagcagcagcttttttctagcaatttttggccacaaacaa 465
      | | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db347  ATTATGACTAGCTATTATTATTATGACTAATTAATTTTAATGTAATTTTAAACCACACTACTA 358
QY   466  ttgatttagtgtattcaaatcacacctaa 494
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db    3537  TTAGTTCAAGTACATTTAAACACCTCTAA 3329

RESULT          2
US-08-107-755A-1/C
Sequence 1, Application US/08107755A
Patent No. 5721352
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Grudl, Michael E.
TITLE OF INVENTION: No. 5713132el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSER: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107,755A
FILING DATE: 19-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,658
FILING DATE: 30-JAN-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEPHONE: (904) 372-5800
SEQUENCE CHARACTERISTICS:
SEQUENCE ID NO: 1:

```

```

; LENGTH: 6768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Amsacta moorei entomopoxvirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (65..1459)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1474..2151
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (2239..2475)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2502..2987
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3080..6091
; NAME/KEY: CDS
; LOCATION: complement (6277..6768)
; US-08-107-755A-1

```

```

Query Match          3.4%; Score 43; DB 1; Length 6768;
Best Local Similarity 47.8%; Pred. No. 0.044;
Matches 187; Conservative 0; Mismatches 200; Indels 4; Gaps 2;

```

```

QY 145 aaccatcagcagtggtgagtgatgagtgatgaactgaactatctgtcttaactcaga 204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2315 AACCATCTACTGATGATGTTGGGAGATTAACCTTTAGAGGTAAATTTACTCTAAA 2256
QY 205 ccaactgataatactcttcgaaacacatgatatctatctatctgtgacttaactaact 264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2255 GCNAAAAAATAAAGACTGTAAATGATTAATTAATTAATTAATTAATTAATTAATTA 2196
QY 265 aaatccaagatattagtagagatgtagtagatgaagtagatgtagatgtagatgtaga 324
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2195 TTATTCATTTTAAATAAATAATGTTAGTATTTACTTAATTAATTAATTAATTAATTA 2136
QY 325 gctaatagttagtagcctaataatgtagtagtagtagtagtagtagtagtagtagtagt 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2135 AAAAGTGTCTGATGATTTTAAATGATTTTAAATGATTTTAAATGATTTTAAATGAT 2076
QY 385 attc-ctgtaataatgtagtagtagtagtagtagtagtagtagtagtagtagtagtagt 442
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2075 CTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2016
QY 443 gcaatttttttgcacaaactaactaagtagtagtagtagtagtagtagtagtagtagtag 502
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2015 ACGGATTTTCTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1958
QY 503 gtgagctctctctagaacttaaccgtagt 533
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1957 GTGATACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1927

```

```

RESULT 3
US-07-991-867B-1/c
; Sequence 1, Application US/07991867B
; Patent No. 5476781
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5476781 Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: David R. Saliwanchik

```

```

; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991,867B
; FILING DATE: 12-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SRO ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8457 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Amsacta moorei entomopoxvirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (65..1459)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1474..2151
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (2239..2475)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2502..2987
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3080..6091
; NAME/KEY: CDS
; LOCATION: complement (6277..6768)
; US-07-991-867B-1

```

```

Query Match          3.4%; Score 43; DB 1; Length 8457;
Best Local Similarity 47.8%; Pred. No. 0.048;
Matches 187; Conservative 0; Mismatches 200; Indels 4; Gaps 2;
QY 145 aaccatcagcagtggtgagtgatgagtgatgaactgaactatctgtcttaactcaga 204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2315 AACCATCTACTGATGATGTTGGGAGATTAACCTTTAGAGGTAAATTTACTCTAAA 2256
QY 205 ccaactgataatactcttcgaaacacatgatatctatctatctgtgtagtagtagtagtag 264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2255 GCNAAAAAATAAAGACTGTAAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2196
QY 265 aaatccaagatattagtagagatgtagtagatgaagtagatgtagatgtagatgtagatg 324

```


Fri Dec 21 09:04:54 2001

us-09-718-754a-1.rni

Page 6

```

1 MEDIUM TYPE: floppy disk
2 COMPUTER: IBM PC compatible
3 OPERATING SYSTEM: PC-DOS/MS-DOS
4 SOFTWARE: ASCII (text)
5 CURRENT APPLICATION DATA:
6 APPLICATION NUMBER: US/08/488,015B
7 FILING DATE: 07-JUN-1995
8 ATTORNEY/AGENT INFORMATION:
9 NAME: Vincent, Matthew P.
10 REGISTRATION NUMBER: 36,709
11 REFERENCE/DOCKET NUMBER: H0V-008,02
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: (617) 832-1000
14 TELEFAX: (617) 832-7000
15 INFORMATION FOR SEQ ID NO: 1:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 4539 base pairs
18 TYPE: nucleic acid
19 STRANDEDNESS: double
20 TOPOLOGY: both
21 MOLECULE TYPE: other nucleic acid
22 FEATURE:
23 NAME/KEY: misc_feature
24 LOCATION: 969..1259
25 OTHER INFORMATION: /product= "E3 exon"
26 FEATURE:
27 NAME/KEY: misc_feature
28 LOCATION: 1290..1559
29 OTHER INFORMATION: /product= "E5 exon"
30
31 US-08-488-015B-1

```

```

Query Match: 3.4%; Score 42; DB 1; Length 4539;
Best Local Similarity 47.5%; P-adj. No. 0.07;
Matches 155; Conservative 0; Mismatches 170; Indels 1; Gaps 1

Oy 20 tatcacatcatcttcacacataatagaagaccataatctaaactaaactaaacac 79
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1822 TATTCATATATATTTATTTATTTATATATATATATATATATATATATAT 1881
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy 80 ttctcacgctacagagagaagagagacataactctgcatagtctttacgtttaacac 139
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1882 TCGATTGGAAATGACATACGATTAATCTCATTAACCATTTGATTAATTTGAAAGCTA 1941
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy 140 ctctgcagccatcagcagtgatgctggttcaactgatactatcctgctcttaac 199
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1942 AGTTAATATTTTACGTATTTTATATATTAACAAATTAACCCCTTAATATTTATTTAA 2001
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy 200 tagcaccacatgataataactcttcgaaacactgtaattatcatcattgctacttaac 259
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2002 TAATTAATAAATATATATATAT -ACCAATATATATATATATATTTATTTATTTATATA 2060
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy 260 atctctaactccagagatttgtagaagcttgctgataagataagagtgattgaatga 319
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2061 ATTAATAATTAATATATATTTATTTATTTATTTGAGATTAGAAATATATATATTTTATGACA 2120
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy 320 ctgagactaatagttaagtgcctcaaa 345
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2121 ATTTCCTTATATTTTATGAGGCTAAAA 2146
    ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-08-814-412-11
; Sequence 11, Application US/08814412
; Patent No. 6150141
; GENERAL INFORMATION:
; APPLICANT: Jarrell Ph.D., Kevin A.
; TITLE OF INVENTION: Intron-Mediated Techniques and Reagents
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: 53 State Street
; CITY: Boston

```

1 STATE: MA
2 COUNTRY: USA
3 ZIP: 02109
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: Patent In Release #1.0, Version #1.30
10
11 CURRENT APPLICATION DATA
12 APPLICATION NUMBER: US/08/814,412
13 FILING DATE: 11-MAR-1997
14
15 CLASSIFICATION: 435
16
17 ATTORNEY/AGENT INFORMATION:
18 NAME: Jarrell Ph.D., Brenda H.
19 REGISTRATION NUMBER: 39,223
20 REFERENCE/DOCKET NUMBER: 0079571-0040
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: 617 248 4000
23
24 INFORMATION FOR SEQ ID NO: 11:
25 SEQUENCE CHARACTERISTICS:
26 LENGTH: 4542 base pairs
27 TYPE: nucleic acid
28 STRANDEDNESS: double
29 TOPOLOGY: circular
30 MOLECULE TYPE: other nucleic acid
31 DESCRIPTION: desc = "plasmid DNA"
32 IMMEDIATE SOURCE:
33 CLONE: PINV1
34
35 US-08-814-412-11

3.4%; Score 42; DB 3; Length 4542;
Query Match Similarity 47.5%; Pred. No. 0.07;
Best Local Similarity 47.5%; Pred. No. 0.07;
Matches 155; Conservative 0; Mismatches 170; Indels 1; Gaps 1.

APPLICATION NUMBER: CH 0016/97
 FILING DATE: 31-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38, 241
 REFERENCE/DOC# NUMBER: PF/5-30306/A/CGCL976
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 186:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 615 base pairs
 TYPE: nucleic acid

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: Patcm Release #1.0, Version #1.30
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/998,416
8  FILING DATE: 24-DEC-1997
9  CLASSIFICATION: 435
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: CH 0016/97
12 FILING DATE: 31-DEC-1996
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Weigs, J. Timothy

```

```

: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8587
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 1137:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 636 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: PAC1692RP
US-08-998-416-1137

```

```

Query Match      3.2%; Score 40.4; DB 4; Length 636;
Best Local Similarity 47.3%; Pred. No. 0.084;
Matches 157; Conservative 0; Mismatches 171; Indels 4; Gaps 1;

QY 160 ttgaaggcttaactgaactaatactgtccttaactaactgaacccaactgaataat 219
   || || || || || || || || || || || || || || || || || || || || ||
DB 45 ttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 220 ctttcgaacatgttatctatctgtgaccttaactaactgaactgaacgaatatta 279
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 105 ttaattatttgatttatttactttatttatttatttatttatttatttatttatttattt
   || || || || || || || || || || || || || || || || || || || || ||

QY 280 gtagagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatg 339
   || || || || || || || || || || || || || || || || || || || || ||
DB 165 ttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 340 cttaaatagcttgtagacatccaacacccatcaattatgatttttttttttttttttt 399
   || || || || || || || || || || || || || || || || || || || || ||
DB 225 tcaaaatttttaatttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 400 gtttaagtagttagttagttagttagttagttagttagttagttagttagttagttag 455
   || || || || || || || || || || || || || || || || || || || || ||
DB 285 ttaatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 456 caactgaacatgtagttagttagttagttagttagttagttagttagttagttagttag 487
   || || || || || || || || || || || || || || || || || || || || ||
DB 345 ttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

```

```

: RESULT 14
: US-08-998-416-1137/c
: Sequence 1137, Application US/08998416
: Patent No. 6239264
: GENERAL INFORMATION:
: APPLICANT: Philippsen, Peter
: APPLICANT: Pohlmann, Rainer
: APPLICANT: Steiner, Sabine
: APPLICANT: Mohr, Christine
: APPLICANT: Wendland, Jürgen
: APPLICANT: Knechtle, Philipp
: APPLICANT: Rebschung, Corinne
: TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
: TITLE OF INVENTION: AND USES THEREOF
: NUMBER OF SEQUENCES: 1152
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6239264artis Corporation
: STREET: 3054 Cornwalis Road
: CITY: Research Triangle Park
: STATE: No. 6239264th Carolina
: COUNTRY: USA
: ZIP: 27709
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30

```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08-998,416
: FILING DATE: 24-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: CH 0016/97
: FILING DATE: 31-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8587
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 1137:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 636 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: PAC1692RP
US-08-998-416-1137

```

```

Query Match      3.2%; Score 40.4; DB 4; Length 636;
Best Local Similarity 43.4%; Pred. No. 0.084;
Matches 235; Conservative 0; Mismatches 306; Indels 1; Gaps 1;

QY 48 aaaccaatttaactaaactaaactaaacttaacacgtagagagagagagagagagagc 107
   || || || || || || || || || || || || || || || || || || || || ||
DB 590 aatttaatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 108 aacttgctatgctttttagcatttaaacacccctcggaacccaagcagcagtgctgtagg 167
   || || || || || || || || || || || || || || || || || || || || ||
DB 530 aaaaatatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 168 tttaactgatactaaactcctgtccttaactaactgaacccaactgataatactctcgaa 227
   || || || || || || || || || || || || || || || || || || || || ||
DB 470 aatttttaatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 228 caccgtgttatctctatgttgaccttaactaactgaacccaactgaactgttgtagaga 286
   || || || || || || || || || || || || || || || || || || || || ||
DB 410 aatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 287 tgttagtagatgaagtgatgttgtagatgaagtagagtagagtagagtagagtagagtag 346
   || || || || || || || || || || || || || || || || || || || || ||
DB 350 ttaatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 347 tagtggagacatccaacacccatcaattatgatttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||
DB 290 aatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 407 gtttagttagttagttagttagttagttagttagttagttagttagttagttagttagttagttag
   || || || || || || || || || || || || || || || || || || || || ||
DB 230 aatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 467 tagttagttagttagttagttagttagttagttagttagttagttagttagttagttagttagttag
   || || || || || || || || || || || || || || || || || || || || ||
DB 170 ttaatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 527 ccgtagtggagagacacatttccatagtgtagttagttagttagttagttagttagttagttagttag
   || || || || || || || || || || || || || || || || || || || || ||
DB 110 aatttttttttttttttttttttttttttttttttttttttttttttttttttttttt
   || || || || || || || || || || || || || || || || || || || || ||

QY 587 ta 588
   || || || || || || || || || || || || || || || || || || || || ||
DB 50 TA 49

```

```

: RESULT 15
: US-08-998-416-288

```

Fri Dec 21 09:04:54 2001

us-09-718-754a-1.rni

Page 10

```

1      Sequence 288, Application US/08998416
2      Patent No. 6239264
3      GENERAL INFORMATION:
4      APPLICANT:  Philippsen, Peter
5      APPLICANT:  Pohlmann, Rainer
6      APPLICANT:  Steiner, Sabine
7      APPLICANT:  Mohr, Christine
8      APPLICANT:  Wendland, Jürgen
9      APPLICANT:  Knechtle, Philipp
10     APPLICANT:  Reibischung, Corinne
11     TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYPII
12     NUMBER OF SEQUENCES:  1152
13     CORRESPONDENCE ADDRESS:
14     ADDRESSEE:  No. 6239264artis Corporation
15     STREET:  3054 Cornwallis Road
16     CITY:  Research Triangle Park
17     STATE:  No. 6239264th Carolina
18     COUNTRY:  USA
19     ZIP:  27709
20     COMPUTER READABLE FORM:
21     MEDIUM TYPE:  floppy disk
22     COMPUTER:  IBM PC compatible
23     OPERATING SYSTEM:  PC-DOS/MS-DOS
24     SOFTWARE:  Patentin Release #1.0, Version #1.30
25     CURRENT APPLICATION DATA:
26     APPLICATION NUMBER:  US/08/998,416
27     FILING DATE:  24-DEC-1997
28     CLASSIFICATION:  435
29     PRIOR APPLICATION DATA:
30     APPLICATION NUMBER:  CH 0016/97
31     FILING DATE:  31-DEC-1996
32     ATTORNEY/AGENT INFORMATION:
33     NAME:  Weigs, J. Timothy
34     REGISTRATION NUMBER:  38, 241
35     REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
36     TELECOMMUNICATION INFORMATION:
37     TELEPHONE:  919-541-8587
38     TELEFAX:  919-541-8689
39     INFORMATION FOR SEQ ID NO: 288:
40     SEQUENCE CHARACTERISTICS:
41     LENGTH:  837 base pairs
42     TYPE:  nucleic acid
43     STRANDEDNESS:  single
44     TOPOLOGY:  linear
45     MOLECULE TYPE:  DNA (genomic)
46     ORIGINAL SOURCE:
47     ORGANISM:  MAG1241RP
48     US-08-998-416-288

```

| | | |
|----|---|--|
| | | 3.2% ; Score 40.4 ; DB 4 ; Length 837 ; |
| | Best Local Similarity | 47.3% ; Score No. 0.094 ; |
| | Matches 157 ; Conservative | 0 ; Mismatches 171 ; Indels 4 ; Gaps 1 ; |
| Qy | 160 ttgatagattcaactgatactatactctgctcttaactagaccacaatgataatg | 219 |
| Db | 45 tttttatttatatttttaactgtatttaatttttaacattattatcatttttataa | 104 |
| Qy | 220 cttcgaacacagcttattctcatgttgaccttaactcaataccaagatctaa | 279 |
| Db | 105 tttaattatttgatttatataactgtattatataattttattttatattttatca | 164 |
| Qy | 280 gtgagagcttagtataagattaaagtgagcttgtaagcactgagctaaatgtag | 359 |
| Db | 165 tttttatttatatttatataattttatatttatgatttttttttttttttttt | 224 |
| Qy | 340 cttaaatctagcttgagacattcaaacacctctcaactatctatagtttttgta | 359 |
| Db | 225 tcataatttttaattgatttttaataatttttttttttttttttttttttttt | 204 |
| Qy | 400 gtttaagctagcttagtattcttaataagcttagctttt---tttaactgaa | 455 |

Dd 285 TAAATTAAGATTACCAATAATTTAATTAATTTTTTAAATGTTAATTAATAATATAATT 344

Qy 456 caactaacacattagtttagtgcattcaaaata 487
| | | | |
Dd 345 TTATATATTATAAGATTATTAATTTAAATA 376

Search completed: December 20, 2001, 12:16:39
Job time: 3882 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 20, 2001, 10:01:26 ; Search time 1550.78 seconds
(without alignments)
8640.804 Million cell updates/sec

Title: US-09-718-754a-1

Sequence: 1 atcgataaagttgagtt.....agagtcgtagcaactagca 1247

Scoring table: IDENTITY_NUC
Gapop 10.0 / Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

EST:*
1: em_estfun:*
2: em_esthm:*
3: em_estlo:*
4: em_estlo:*
5: em_estlo:*
6: em_estlo:*
7: em_estlo:*
8: em_estlo:*
9: em_estlo:*
10: em_estlo:*
11: em_estlo:*
12: em_estlo:*
13: em_estlo:*
14: em_estlo:*
15: em_estlo:*
16: em_estlo:*
17: em_estlo:*
18: em_estlo:*
19: em_estlo:*
20: em_estlo:*
21: em_estlo:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 126.6 | 10.2 | 365 | 10 | AI619399 486093611 |
| 2 | 126.6 | 10.2 | 407 | 10 | AI657360 486093611 |
| 3 | 118.6 | 9.5 | 687 | 10 | AM066683 683006010 |
| 4 | 93.8 | 7.5 | 616 | 10 | AM147120 707013910 |
| 5 | 93.8 | 7.5 | 620 | 11 | BF727815 100005301 |
| 6 | 86.8 | 7.0 | 497 | 10 | AM258129 687068610 |
| 7 | 75 | 6.1 | 474 | 13 | BH130093 |
| 8 | 75 | 6.1 | 1101 | 13 | CNS0057L |
| 9 | 67.8 | 5.4 | 412 | 10 | BE510052 |
| 10 | 65.6 | 5.3 | 1101 | 13 | CNS00172 |
| 11 | 64.6 | 5.2 | 438 | 13 | AQ844669 an43f03 J |
| 12 | 63.6 | 5.1 | 1101 | 13 | CNS003BB |

| | | | | | | |
|------|------|-----|------|----|-----------|----------------------|
| C 13 | 63.4 | 5.1 | 1101 | 13 | CNS0057L | AL069706 Drosophila |
| C 14 | 63.2 | 5.1 | 928 | 13 | CNS000DKY | AL071865 Drosophila |
| C 15 | 61.6 | 4.9 | 928 | 13 | CNS000DKY | AL071865 Drosophila |
| C 16 | 61.4 | 4.9 | 192 | 10 | AI612358 | AL612358 486079H11 |
| C 17 | 61.4 | 4.9 | 1125 | 10 | AL547503 | AL547503 486079H11 |
| C 18 | 61.2 | 4.9 | 1027 | 13 | CNS02PT50 | AL212733 Tetradodon |
| C 19 | 60.8 | 4.9 | 550 | 10 | AI600397 | AI600397 486071A03 |
| C 20 | 60.8 | 4.9 | 1101 | 13 | CNS0039C | AL175696 Tetradodon |
| C 21 | 59.6 | 4.8 | 496 | 13 | BH127174 | BH127174 G-1019-1 |
| C 22 | 57.4 | 4.6 | 1101 | 13 | CNS0039C | AL063921 Drosophila |
| C 23 | 56.4 | 4.5 | 836 | 13 | CNS01100 | AL069440 Drosophila |
| C 24 | 56.4 | 4.5 | 1101 | 13 | CNS000E07 | AL069440 Drosophila |
| C 25 | 56.2 | 4.5 | 894 | 13 | BH136681 | BH136681 ENTOR72TF |
| C 26 | 55.4 | 4.4 | 773 | 13 | BG584434 | BG584434 1024040A0 |
| C 27 | 55.2 | 4.4 | 758 | 10 | AV756150 | AV756150 AV756150 |
| C 28 | 55.2 | 4.4 | 1001 | 13 | CNS0155H | AL105023 Drosophila |
| C 29 | 54.8 | 4.4 | 1190 | 13 | CNS020N7 | AL206908 Tetradodon |
| C 30 | 54.6 | 4.4 | 994 | 13 | CNS04NDJ | AL298972 Tetradodon |
| C 31 | 54.6 | 4.4 | 1092 | 13 | CNS020K7 | AL175696 Tetradodon |
| C 32 | 54.4 | 4.4 | 1101 | 13 | CNS0021J | AL1061936 Drosophila |
| C 33 | 54.4 | 4.4 | 1225 | 13 | CNS0161D | AL1061936 Drosophila |
| C 34 | 54.2 | 4.3 | 1796 | 10 | AV704385 | AV704385 AV704385 |
| C 35 | 53.8 | 4.3 | 1086 | 13 | CNS0017K | AL069682 Drosophila |
| C 36 | 53.6 | 4.3 | 1101 | 13 | CNS000LT2 | AL078714 Drosophila |
| C 37 | 53.6 | 4.3 | 1161 | 13 | CNS0073Y8 | AL078714 Drosophila |
| C 38 | 53.6 | 4.3 | 1201 | 13 | CNS0167M | AL0633529 Drosophila |
| C 39 | 53.4 | 4.3 | 876 | 13 | BF296544 | BF296544 038PB212 |
| C 40 | 53.2 | 4.3 | 619 | 11 | CNS014P0 | AL104456 Drosophila |
| C 41 | 53.2 | 4.3 | 987 | 13 | CNS014P0 | AL104456 Drosophila |
| C 42 | 53 | 4.3 | 994 | 13 | CNS04NDJ | AL298972 Tetradodon |
| C 43 | 52.8 | 4.2 | 1085 | 13 | CNS016YR | AL107373 Drosophila |
| C 44 | 52.8 | 4.2 | 1101 | 13 | CNS0022U | AL0977152 Drosophila |
| C 45 | 52.6 | 4.2 | 139 | 13 | AZ916538 | AZ916538 PstL7_c4 |

ALIGNMENTS

RESULT 1
AI619399/c 365 bp mRNA EST 21-APR-1999
LOCUS 486093611.x1 486 - leaf primordia CDNA library from Hake lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION AI619399
VERSION AI619399.1 GI:4628525
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade: Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 365)
AUTHORS Walbot V.
TITLE maize ESTs from various CDNA libraries sequenced at Stanford
COMMENT
JOURNAL Unpublished (1999)
CONTACT: Walbot V
DEPARTMENT: Department of Biological Sciences
STANFORD UNIVERSITY
855 California Ave, Palo Alto, CA 94304, USA
TEL: 650 723 2227
FAX: 650 725 8221
EMAIL: walbot@stanford.edu
PLATE: 486093 row: G column: 11.
LOCATION/Qualifiers
1. 365
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_id="486 - leaf primordia CDNA library from Hake
lab"
/tissue="leaf primordia"
/dev_stage="P7-P11 leaf"

Page 2

| | | | | |
|---------------------------|--------|--------------------|-----------|-------------|
| Query Match | 10.2%; | Score 126.6; | DB 10; | Length 407; |
| Best Local Similarity | 79.4%; | Pred. NO. 2.3e-15; | | |
| Matches 166; Conservative | 0; | Mismatches 34; | Indels 9; | Gaps 1 |

| | | | | |
|-----------------------|---|-------------------|-----------|-------------|
| Query Match | 9.5% | Score 118.6; | DB 10; | Length 687; |
| Best Local Similarity | 93.2%; | Ped. No. 7.9e-14; | | |
| Matches 124; | Conservative 0; | Mismatches 9; | Indels 0; | Gaps |
| QY 1115 | cgagagtagtgcgcatgctgtaacgtcagctgtgaagctttttgcagagagctgcgtctgcaaatag | 1174 | | |
| | | | | |
| DB 31 | CGGAGACTACTCTCTGTGTACCGCTACTGACTGCTTTTGGCAGAGCCTCTCTGTAATATAG | 90 | | |
| QY 1175 | tagcccttcacaacagcgcgcgaaggggggagagatgacgtlcaagcttagcagagagatgc | 1234 | | |
| | | | | |
| DB 91 | TAGCCCTTCCACAGACGAGCGGAGCGGGAGAGAGTATGTGCTAGCTTACGAGAGAGTGC | 150 | | |

Fri Dec 21 09:04:55 2001

us-09-718-754a-1.rst

Page 3

QY 1235 gtagcaactagca 1247
|||||
Db 151 GTACCACTAGCA 163

RESULT 4
LOCUS AM147120 616 bp mRNA
DEFINITION 707013F10.x1.707 - Mixed adult tissues from Walbot lab (SK) Zea
ACCESSION AM147120
VERSION AM147120.1 GI:6195016
KEYWORDS EST
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 616)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707013 row: F column: 10.
Location/Qualifiers
1. 616
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10a"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EORI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT 201 a 107 c 88 g 218 t 2 others

ORIGIN

Query Match 7.5%; Score 93.8; DB 10; Length 616;
Best Local Similarity 71.6%; Pred. No. 6.5e-09;
Matches 141; Conservative 0; Mismatches 47; Indels 9; Gaps 1;

QY 308 ttttgaatgacactagctaaatagctaaatagctgagacatccaacac 367
|||||
Db 135 ttttgaatgacactagctaaatagctaaatagctgagacatccaacac 194

QY 368 ccta-----tcaattatagttatctttagtaaatagttcaatagtttagtca 418
|||||
Db 195 gctagatgataattatgactagctaaatagctaaatagctgagacatccaacac 254

QY 419 ttataagcactagctttttactagcaatttttagcacaacaatagtttagt 478
|||||
Db 255 tttgttgaactgataatgtccgctagcgatTTTTTAACCACTAATATATCTGTGTC 314

QY 479 attcaaatccccaag 495
|||||
Db 315 ATTCAAACTCTATTAG 331

RESULT 5

BF727815
LOCUS BF727815 620 bp mRNA EST
DEFINITION 1000053C11.x2.1000 - Unigene 1 from Maize Genome Project Zea mays
ACCESSION BF727815
VERSION BF727815.1 GI:12045676
KEYWORDS EST
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 620)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 100053 row: C column: 11.
Location/Qualifiers
1. 620
/organism="Zea mays"
/db_xref="dbEST:707013F10.x1"
/db_xref="taxon:4577"
/note="This library represents the unique ESTs found in
the first round of EST sequencing at Stanford University
for the maize genome project. Sequences are present from
libraries 486, 487, 496, 603, 605, 614, 618, 660, 683
687, 707, and 945. Contigs were assembled using TIGR's
CAP program and a representative EST from each contig was
selected."

BASE COUNT 200 a 107 c 86 g 227 t

ORIGIN

Query Match 7.5%; Score 93.8; DB 11; Length 620;
Best Local Similarity 71.6%; Pred. No. 6.5e-09;
Matches 141; Conservative 0; Mismatches 47; Indels 9; Gaps 1;

QY 308 ttttgaatgacactagctaaatagctaaatagctgagacatccaacac 367
|||||
Db 145 ttttgaatgacactagctaaatagctaaatagctgagacatccaacac 204

QY 368 ccta-----tcaattatagttatctttagtaaatagtttagtca 418
|||||
Db 205 gctagatgataattatgactagctaaatagctaaatagctgagacatccaacac 264

QY 419 ttataagcactagctttttactagcaatttttagcacaacaatagtttagt 478
|||||
Db 265 tttgttgaactgataatgtccgctagcgatTTTTTAACCACTAATATATCTGTGTC 324

QY 479 attcaaatccccaag 495
|||||
Db 325 ATTCAAACTCTATTAG 341

RESULT 6
LOCUS AM258129 497 bp mRNA EST
DEFINITION 687068G10.y1.687 - Early embryo from Delaware Zea mays cDNA, mRNA
ACCESSION AM258129
VERSION AM258129.1 GI:6626597
KEYWORDS EST
SOURCE Zea mays.

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; on bw sp, the same strain used for the BDP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers
 1. 1101
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_lib="RPCI-98"
 /clone="BACR48P19"
 /note="end : rev3"
 BASE COUNT 469 a 6 c 69 g 151 t 406 others
 ORIGIN

Query Match

Best Local Similarity 5.38; Score 65.6; DB 13; Length 1101;
 Matches 185; Conservative 114; Mismatches 303; Indels 0; Gaps 0;

```

OY 6 acataagttgattacataatccattcaactaataagaacaaatttaataa 65
DB 257 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 316
OY 66 aactaaaaaaccttaccgtagagagagagagagagagagagagagagag 125
DB 317 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 376
OY 126 tgcatttaaacacccttcagacatcagcagtggtgtagtgtagttagttagt 185
DB 377 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 436
OY 186 ctgtcttcttactagcaccacgtatataatcttcgagacagctgattatcat 245
DB 437 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 496
OY 246 gttagctttaaataactaataccagatattagtagagattgtagttagttagt 305
DB 497 TTTTATTTTWTWTTTAAADDBWAAWMTTWTWTTTAAATTTTWTWTTWK 556
OY 306 gattgttagaagcctcagcagcagcagcagcagcagcagcagcagcagcagc 365
DB 557 ATTTDTTAKAKYAAKTAAMAAWTTTAAATTTTAAATTTTAAATTTTAAATTT 616
OY 366 accctataattatagttattttagtaattagtttagttagttagttagttagt 425
DB 617 WTTWTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 676
OY 426 gctagcttttcttagcagcagcagcagcagcagcagcagcagcagcagcagc 485
DB 677 AAAAAAAAAAAMWTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAA 736
OY 486 taaccttaagccgttaagtagcgtcttcctcagaatccttaacggtatgagagaa 545
DB 737 WAAADKKDKKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 796
OY 546 tctcaatagtgtagttagttagttagttagttagttagttagttagttagttagt 605
DB 797 KKAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 856
OY 606 tt 607
DB 857 TT 858

```

RESULT 11
 A0844669 438 bp DNA GSS 03-JUN-1999
 LOCUS A0844669/c
 DEFINITION an4f03 JM107 filtered library zea mays genomic. DNA sequence.
 ACCESSION A0844669

VERSION A0844669.1 GI:6203157

KEYWORDS

GSS.

SOURCE

zea mays.
 zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 438)
 Rabl, J., McCord, W.R., Dedhia, N., Jordan, C., Parnell, L.D.,

AUTHORS

Stein, L., McCord, W.R., and Martienssen, R.A.

TITLE

Differential methylation of genes and retrotransposons allows

JOURNAL

shotgun sequencing of the maize genome
 Nat. Genet. 23, 305-308 (1999)

COMMENT

Contact: Martienssen RA
 Cold Spring Harbor Laboratory
 1 Bungtown Rd., Cold Spring Harbor, NY 11724, USA
 Tel: 516 367 8322
 Fax: 516 367 8369
 Email: martienss@cshl.org
 Seq primer: forward
 Class: shotgun.

FEATURES

Location/Qualifiers

1. 438
 /organism="Zea mays"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone_lib="JM107 filtered library"
 /note="Organ: Immature ears; Vector: M13; Site: 1; Xba I;
 DNA prepared from purified nuclei was digested with the
 methylation insensitive enzyme Spe I, size fractionated to
 enrich for the 0.5 to 4 kbp fraction, ligated into Xba I
 digested M13 vector and electroporated into E.coli JM107."

BASE COUNT 63 a 138 c 128 g 108 t 1 others
 ORIGIN

Query Match

Best Local Similarity 5.28; Score 64.6; DB 13; Length 438;
 Matches 73; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```

OY 294 atagattaaagtgatttgtagcactagagcagcagcagcagcagcagcagcagc 353
DB 91 ATACTTAAAGAGAGTGTGATGACTAGAGCTAATAGTGTAGTGTAAATTAAGTACT 32

```

```

OY 354 agacattcaaacacccatcaattatt 380
DB 31 AGACATCAACAACACTAGCTAATATT 5

```

RESULT 12
 A0844669 1101 bp DNA GSS 03-JUN-1999
 LOCUS A0844669/c
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #

BACR08008 of RPCI-98 library from Drosophila melanogaster (fruit fly) genomic survey sequence.

ACCESSION A0844669
 VERSION A0844669.1 GI:4941845

KEYWORDS

GSS.

SOURCE

fruit fly.
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1101)
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage ;
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

TITLES

Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).

AUTHORS

The BDGP is constructing a physical map of the Drosophila

COMMENT

QY 353 gagacattcaaacaccctatcaattatagttatttttagtaaatagtaagtagt 412
Db 626 AATATTTTMTTAAATTAATAATTAATAATTAATAATTAATAATTAATAAT 567
QY 413 tagttatttaagctagcttttttactagcaatttttagccaactacaa:tagtt 472
Db 566 AAAAAAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAAT 507

Search completed: December 20, 2001, 11:40:55
Job time: 5969 sec